What is claimed is:

1   1.   A method of instantiating a data placement heuristic for a distributed storage
2   system comprising the steps of:
3   receiving heuristic parameters; and
4   running an algorithm on a node of the distributed storage system which
5   instantiates a particular data placement heuristic selected from a range of
6   data placement heuristics according to the heuristic parameters.

1   2.   A method of instantiating a data placement heuristic for a distributed storage
2   system comprising the steps of:
3   receiving heuristic parameters which comprise a cost function, a
4   placement constraint, a metric scope, an approximation technique, and an
5   evaluation interval; and
6   running an algorithm on a node of the distributed storage system which
7   instantiates a particular data placement heuristic selected from a range of
8   data placement heuristics according to the heuristic parameters.

1   3.   The method of claim 2 wherein the heuristic parameters further comprise at
2   least one additional placement constraint.

1   4.   The method of claim 2 wherein the heuristic parameters further comprise a
2   routing knowledge parameter.

1   5.   The method of claim 2 wherein the heuristic parameters further comprise an
2   activity history parameter.

1   6.   The method of claim 2 wherein the metric scope comprises a node scope,
2   which ranges from a single node of the distributed storage system to all nodes of
3   the distributed storage system.

1   7.   The method of claim 6 wherein the metric scope further comprises a client
2   scope, which ranges from local clients accessing the single node to all clients
3   accessing all of the nodes.

1   8.   The method of claim 6 wherein the metric scope further comprises a data
2   object scope, which ranges from local data objects stored on the single node to all
3   data objects stored on all of the nodes.

1   9.   The method of claim 2 wherein the heuristic parameters further comprise at
2   least one additional approximation technique.

1   10.   The method of claim 2 wherein the algorithm comprises a determination of
2   costs for nodes and data objects within the metric scope according to the cost
3   function.

1   11.   The method of claim 10 wherein the algorithm further comprises a sort of the
2   costs.

1   12.   The method of claim 11 wherein the approximation technique comprises a
2   ranking technique.

1   13.   The method of claim 12 wherein the ranking technique comprises a greedy
2   technique.

1   14.   The method of claim 11 wherein the approximation technique comprises a
2   threshold technique.

1   15.   The method of claim 2 wherein the approximation technique comprises a
2   random technique.

1   16.   The method of claim 2 wherein the approximation techniques comprises an
2   improvement technique.

1   17.   The method of claim 2 wherein the approximation technique comprises a
2   hierarchical technique.

1   18.   The method of claim 2 wherein the approximation technique comprises a

2   multiphase technique.

1   19.   A method of instantiating a data placement heuristic for a distributed storage

2   system comprising the steps of:

3   receiving heuristic parameters which comprise a cost function, a

4   placement constraint, a metric scope, an approximation technique, and an

5   evaluation interval; and

6   for each evaluation interval within an execution of the data placement

7   heuristic, performing the steps of:

8   for data objects and nodes within the metric scope, determining

9   a cost array which comprises cost elements determined from the

10   cost function;

11   if the approximation technique comprises a ranking technique,

12   placing the data objects according to the placement constraint in

13   order of rank within a sorted cost array;

14   if the approximation technique comprises a greedy ranking

15   technique, placing the data objects according to the placement

16   constraint in order of the rank within the sorted cost array, and

17   recomputing and resorting the cost array for remaining cost

18   elements after each placement decision; and

19   if the approximation technique comprises a threshold

20   technique, eliminating the cost elements of the cost array which

21   violate a threshold, thereby forming a threshold limited cost array,

22   and placing the data objects according to the placement constraint

23   in order of rank within a sorted version of the threshold limited cost

24   array.

1   20.   The method of claim 19 wherein the cost function comprises a rate of read

2   accesses by a client for a particular data object.

1   21.   The method of claim 19 wherein the cost function comprises a rate of write

2   accesses by a client for a particular data object.

1   22.   The method of claim 19 wherein the cost function comprises a measure of a

2   distance between a particular client and a particular node.

1   23.   The method of claim 19 wherein the cost function comprises a storage cost for

2   storing a particular data object on a particular node.

1   24.   The method of claim 19 wherein the cost function comprises an access time

2   comprising a timestamp indicating a most recent time when a particular data

3   object was accessed at a particular node.

1   25.   The method of claim 19 wherein the cost function comprises a load time

2   comprising a timestamp indicating a time when a particular data object was stored

3   on a particular node.

1   26.   The method of claim 19 wherein the cost function comprises a hit ratio of

2   transparent en route caches along a path from a client to a node.

1   27.   The method of claim 19 wherein the placement constraint comprises a storage

2   capacity constraint, which places an upper limit on a storage capacity for a node.

1   28.   The method of claim 19 wherein the placement constraint comprises a load

2   capacity constraint, which places an upper limit on a rate of requests for a node.

1   29.   The method of claim 19 wherein the placement constraint comprises a node

2   bandwidth capacity, which places an upper limit data transmission rate for a node.

1   30.   The method of claim 19 wherein the placement constraint comprises a link

2   capacity constraint, which places an upper limit on bandwidth between first and

3   second nodes.

1   31.   The method of claim 19 wherein the placement constraint comprises replica

2   constraint, which places an upper limit on a number of replicas that can be placed

3   on nodes of the distributed storage system.

1   32.   The method of claim 19 wherein the placement constraint comprises a delay

2   constraint, which places an upper limit on a response time for requests within the

3   distributed storage system.

1   33.   The method of claim 19 wherein the placement constraint comprises an

2   availability constraint, which places a lower limit on availability of data objects

3   within the distributed storage system.

1   34.   The method of claim 19 wherein the approximation technique further

2   comprises a hierarchical technique.

1   35.   The method of claim 19 wherein the approximation technique further

2   comprises a multiphase technique.

1   36.   The method of claim 35 wherein the multiphase technique comprises an

2   improvement technique.

1   37.   A computer readable memory comprising computer code for implementing a

2   method of instantiating a data placement heuristic for a distributed storage system,

3   the method of instantiating the data placement heuristic comprising the steps of:

4   receiving heuristic parameters; and

5   running an algorithm on a node of the distributed storage system which

6   instantiates a particular data placement heuristic selected from a range of

7   data placement heuristics according to the heuristic parameters.

1   38.   A computer readable memory comprising computer code for implementing a

2   method of instantiating a data placement heuristic for a distributed storage system,

3   the method of instantiating the data placement heuristic comprising the steps of:

4   receiving heuristic parameters which comprise a cost function, a

5   placement constraint, a metric scope, an approximation technique, and an

6   evaluation interval; and

7   running an algorithm on a node of the distributed storage system which

8   instantiates a particular data placement heuristic selected from a range of

9   data placement heuristics according to the heuristic parameters.

1   39.   A computer readable memory comprising computer code for implementing a

2   method of instantiating a data placement heuristic for a distributed storage system,

3   the method of instantiating the data placement heuristic comprising the steps of:

4   receiving heuristic parameters which comprise a cost function, a

5   placement constraint, a metric scope, an approximation technique, and an

6   evaluation interval; and

7   for each evaluation interval within an execution of the data placement

8   heuristic, performing the steps of:

9   for data objects and nodes within the metric scope, determining

10   a cost array which comprises cost elements determined from the

11   cost function;

12   if the approximation technique comprises a ranking technique,

13   placing the data objects according to the placement constraint in

14   order of rank within a sorted cost array;

15   if the approximation technique comprises a greedy ranking

16   technique, placing the data objects according to the placement

17   constraint in order of the rank within the sorted cost array, and

18   recomputing and resorting the cost array for remaining cost

19   elements after each placement decision; and

20   if the approximation technique comprises a threshold

21   technique, eliminating the cost elements of the cost array which

22   violate a threshold, thereby forming a threshold limited cost array,

23   and placing the data objects according to the placement constraint

24   in order of rank within a sorted version of the threshold limited cost

25   array.